Deep learning based event reconstruction for the Limadou High-Energy Particle Detector

S. Bartocci[1], R. Battiston[2,3], F. Benotto[4], S. Beolè[4,5], W. J. Burger[3,6], D. Campana[7], G. Castellini[8], P. Cipollone[1], S. Coli[4], L. Conti[1,9], A. Contin[10,11], M. Cristoforetti[12], L. de Cilladi[4,5], C. De Donato[1], C. De Santis[1], F. M. Follega[2,3,*], G. Gebbia[2,3], R. Iuppa[2,3,†], M. Lolli[11], N. Marcelli[1,13], M. Martucci[1,13], G. Masciantonio[1], M. Mergè[1,13], M. Mese[7,14], C. Neubüser[3], F. Nozzoli[3], A. Oliva[11], G. Osteria[7], L. Pacini[15], F. Palma[1,13], F. Palmonari[10,11], A. Parmentier[16], F. Perfetto[7], P. Picozza[1,13], M. Piersanti[16], M. Pozzato[11], E. Ricci[2,3], M. Ricci[17], S. B. Ricciarini[8], Z. Sahnoun[11], V. Scotti[7,14], A. Sotgiu[1,13], R. Sparvoli[1,13], V. Vitale[1], S. Zoffoli[18], and P. Zuccon[2,3]

(Limadou HEPD collaboration)

[1]INFN-Sezione di Roma "Tor Vergata", V. della Ricerca Scientifica 1, I-00133 Rome, Italy
[2]University of Trento, V. Sommarive 14, I-38123 Povo (Trento), Italy
[3]INFN-TIFPA, V. Sommarive 14, I-38123 Povo (Trento), Italy
[4]INFN-Sezione di Torino, Via P. Giuria 1, I-10125 Torino, Italy
[5]University of Torino, Via P. Giuria 1, I-10125 Torino, Italy
[6]Centro Fermi, V. Panisperna 89a, I-00184 Rome, Italy
[7]INFN-Sezione di Napoli, V. Cintia, I-80126 Naples, Italy
[8]IFAC-CNR, V. Madonna del Piano 10, I-50019 Sesto Fiorentino (Florence), Italy
[9]Uninettuno University, C.so V. Emanuele II 39, I-00186 Rome, Italy
[10]University of Bologna, V.le C. Berti Pichat 6/2, I-40127 Bologna, Italy
[11]INFN-Sezione di Bologna, V.le C. Berti Pichat 6/2, I-40127 Bologna, Italy
[12]Fondazione Bruno Kessler, V. Sommarive 18, I-38123 Povo (Trento), Italy
[13]University of Rome "Tor Vergata", V. della Ricerca Scientifica 1, I-00133 Rome, Italy
[14]University of Naples "Federico II", V. Cintia 21, I-80126 Naples, Italy
[15]INFN-Sezione di Firenze, V. Sansone 1, I-50019 Sesto Fiorentino (Florence), Italy
[16]INAF-IAPS, V. Fosso del Cavaliere 100, I-00133 Rome, Italy
[17]INFN-LNF, V. E. Fermi 54, I-00044 Frascati (Rome), Italy
[18]Italian Space Agency, V. del Politecnico, I-00133 Rome, Italy

Deep learning algorithms have gained importance in particle physics in the last few years. They have been shown to outperform traditional strategies in particle identification, tracking and energy reconstruction in the most modern high-energy physics experiments. The attractive feature of these techniques is their ability to model large dimensionality inputs and catch nontrivial correlations among the variables, which could be hidden or not easy to model. This paper focuses on the application of deep neural networks to the event reconstruction of the Limadou High-Energy Particle Detector on board the China Seismo-Electromagnetic Satellite. The core of the reconstruction chain is a set of fully connected neural networks that reconstructs the nature, the arrival direction and the kinetic energy of incoming electrons and protons, starting from the signals recorded in the detector. These networks are trained on a dedicated Monte Carlo simulation as representative as possible of real data. We describe the simulation, architecture and methodology adopted to design and train the networks, and finally report on the performance measured on simulated and flight data.

I. INTRODUCTION

The latest particle physics and astrophysics experiments have been increasingly including deep learning (DL) based models in their event reconstruction procedures. Examples are available in particle identification and tracking [1], as for offline improvement in the estimation of physics observables [2]. This approach has the main advantage that correlations between variables are discovered and exploited automatically by these algorithms via the minimization of a cost function, called loss function, during the learning procedure. This can help not only because, in principle, physical intuition can miss some of these

*francesco.follega@unitn.it
†roberto.iuppa@unitn.it

correlations, but also because the machine can automatically take into account detector asymmetries and data inhomogeneities to improve its prediction. Moreover, this approach is particularly powerful when dealing with particle detectors. Indeed these instruments output multiple electrical signals, which can be treated as low-level features. It has been proven that deep neural networks (DNN) work remarkably well on low level features [3]. Therefore, we can directly input the calibrated signals to this algorithm and let it learn how to use correlations to reconstruct interesting event features. A clear example in this direction is reported in this work, where the event reconstruction chain for particles traversing the High-Energy Particle Detector (HEPD) on board the China Seismo Electromagnetic Satellite (CSES-01) is fully based on Fully Connected Neural Networks (FCNNs). The networks take the signals from the photomultiplier tubes as the input, and return the type of particle, the arrival direction in the local frame and the energy as the output.

The article is organized as follows. In Sec. II, the Limadou HEPD is described, highlighting how the information from each sub-detector enters the event reconstruction chain. In Sec. III, details about the Monte Carlo (MC) simulation are provided, paying attention to both training and validation of FCNNs, and to observable features for the comparison with data. Section IV explains the architecture of the networks and the training procedure. Sections V and VI provide information on the performance achieved on MC and data events respectively. Conclusions and perspectives follow.

II. THE LIMADOU HIGH-ENERGY PARTICLE DETECTOR

The reference case of this study is the Limadou HEPD instrument [4], which represents the main Italian contribution to the CSES-01 mission [5,6]. It is a particle detector orbiting around the Earth on board the CSES satellite, and it is designed to: measure fluxes of energetic charged particles trapped in the Van Allen Belts [7], monitor impulsive solar activity [8], and extend the lower part of the cosmic ray spectrum [9]. Moreover, one of the main scientific goals of the mission is to look for possible correlations between particle bursts, i.e., anomalous high fluxes measured during flight in a given time interval, and medium-strong seismic events [10,11]. Indeed several studies have reported observations that electromagnetic emissions induced by earthquakes are able to propagate toward the magnetosphere [12,13]. Indeed these events may produce low frequency electromagnetic waves that propagate toward the magnetosphere [14,15] and interact with charged particles in the Van Allen Belts causing pitch angle diffusion of the trapped particles, inducing their precipitation [16].

The Limadou HEPD has been designed to detect electrons with kinetic energy 3–100 MeV and protons in the range 30–200 MeV. A technical drawing of the HEPD is
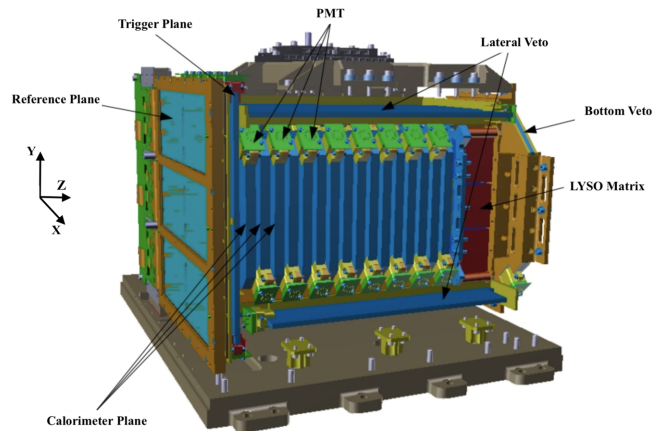

FIG. 1.   Limadou HEPD CAD rendering.

shown in Fig. 1 and it is composed of different subsystems. For the purposes of this work, the signals of the following subdetectors are used:

  (i) Trigger: a plane made of plastic scintillators, segmented into six paddles ($20 \times 3 \times 0.5$ cm^3), each one read by two photo multipliers tubes (PMTs);

 (ii) Upper calorimeter: sixteen $15 \times 15 \times 1$ cm^3 planes of plastic scintillator, each one read out by two PMTs. The two PMTs are evenly distributed on the edges of each plane in such a way that any two subsequent planes have their two PMTs placed along opposite diagonals;

(iii) Lower calorimeter: a 3×3 matrix ($5 \times 5 \times 4$ cm^3) of LYSO (Lutetium-Yttrium Oxyorthosilicate) inorganic scintillator crystals, each one read out by one PMT;

(iv) Veto: five planes of plastic scintillators (0.5 cm thick) each one read out by 2 PMTs at the edges, surrounding the detector.

Figure 1 also shows the system of local coordinates used in this work. The polar angle θ is measured from the z axis, whereas the azimuthal angle ϕ is measured on the x–y plane, from the x axis toward the y axis. A plane placed at $\Delta z = 20$ mm upstream of the trigger plane has been chosen as reference plane for the entrance of particles in the HEPD simulation volume.

III. GEANT4 SIMULATION AND DIGITIZATION

The ideal case for training deep learning (DL) algorithms is to have a large set of data from the experiment that can be used as ground truth. In many circumstances, like the one described in this work, these data are not available or are not sufficiently dense to suitably cover all the phase space explored during real data taking. For this reason, the Limadou collaboration devoted time and resources to the production of a suitable simulation sample, which has been used as the input dataset to train DL algorithms.

The networks presented in this article were trained on a sample of protons and a sample of electrons. Their interactions with materials and the response of the detector have been simulated using GEANT4 [17]. An example of a 200 MeV proton interaction in HEPD is shown in Fig. 2.

The simulated detector is based on the CAD drawing of the instrument. Both mechanical structures and sensitive materials have been simulated to correctly reproduce the detector response. Also the readout chain has been simulated, including the optical photon generation from the plastic scintillators and LYSO scintillating crystals, as well as the conversion of light to electric signal as for the real instrument. The GEANT4 simulation relies on the physics lists accounting for all physics processes accessible at the simulated energy range. In these lists electromagnetic and hadronic interaction, as well as the interaction with heavy nuclei and ions, are included [17].

The events have been generated to suitably fill the parameter space of the HEPD data acquisition in beam tests and during flight. Multiple interactions with HEPD for protons and electrons with the following features have been simulated:

 (i) Kinetic energy: $10 < E^{\mathrm{Kin}}_{\mathrm{prot}} < 1000$ MeV and $1 < E^{\mathrm{Kin}}_{\mathrm{elec}} < 200$ MeV;
 (ii) Polar angle: $0° < \theta < 90°$, sampled uniformly in $\cos^2 \theta$ in order to guarantee an isotropic flux;
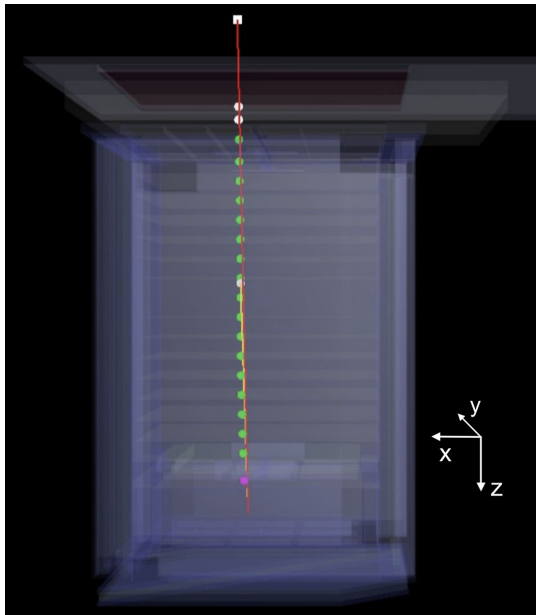 (iii) Azimuthal angle: $-180° < \phi < 180°$.

FIG. 2. Event display of a nearly-vertical 200 MeV proton. Sensitive volumes are overlaid in blue, support structures and auxiliary volumes in gray. The red track represents the primary proton, while the white square on top is the generation point. The points represent the energy deposits from particles to materials. Colors represent the amount deposited energy: white corresponds to ~0.1 MeV, green to ~1 MeV, and purple to ~100 MeV.

 (iv) particles have been generated uniformly from a 40×40 cm^2 window (parallel to the x-y plane) right above the reference plane.

Electrons (protons) with energy higher than 100 (200) MeV do not stop their trajectory inside the detector, so that the energy release does not amount to the total energy. These particles are referred to as "noncontained" hereafter.

Signals are simulated from the energy deposited by primary and secondary particles in scintillating materials. The energy deposit determines the amount of scintillation (about 10^4 photons/MeV in HEPD plastic scintillators). After the generation, photons propagate in the material and are followed until they either escape the simulation volume or hit a PMT. In the latter case, the quantum efficiency for converting photons to photoelectrons is accounted for and the signal is generated accordingly. The threshold for trigger and readout electronics has been implemented as $E_{\mathrm{dep}} \geq E_{\mathrm{m.i.p.}}/4$, where E_{dep} represents the total energy deposited in the sensitive volume and $E_{\mathrm{m.i.p.}}$ the modal energy released by a minimum ionizing particle (m.i.p.) in the same volume. The following strategy has been envisaged for the signal digitization: photoelectrons are converted to ADC units by comparing PMT response distributions from beam tests and simulations [18].

Simulated events undergo preselection to reproduce online trigger masks controlling the HEPD data acquisition. Trigger masks always involve signals from the trigger plane, logically combined with signals from plastic scintillator planes. This work has been carried out after the following preselection:

$$T \text{ AND } P_1 \text{ AND } P_2 \text{ AND NOT (LatVeto)} \qquad (1)$$

where T indicates a signal from the trigger plane, P_1 and P_2 a signal from the first two plastic scintillators, respectively, and LatVeto a signal from the lateral veto. Preselection (1) reproduces the trigger mask mostly used for HEPD since the launch of CSES in 2018.

In Fig. 3 the generated distributions of polar angle and energy are shown for electrons before and after the preselection (similar behavior has been observed for protons).

After selection (1), the energy distribution shows a clear threshold effect due to the minimum energy necessary to electrons to reach the scintillator plane P_2, as well as the fractional incidence of high-energy events diminishes due to inclined electrons rejected under the NOT(LatVeto) condition. The same condition sculpts the θ distribution as well, favoring quasi on-axis events over inclined ones. The number of generated protons amounts to $N^{\mathrm{prot}}_{\mathrm{gen}} = 5.263 \times 10^6$, whereas $N^{\mathrm{prot}}_{\mathrm{pres}} = 5.060 \times 10^5$ pass the preselection (1). For electrons, the numbers are $N^{\mathrm{elec}}_{\mathrm{gen}} = 6.921 \times 10^6$ and $N^{\mathrm{elec}}_{\mathrm{pres}} = 6.952 \times 10^5$, respectively.

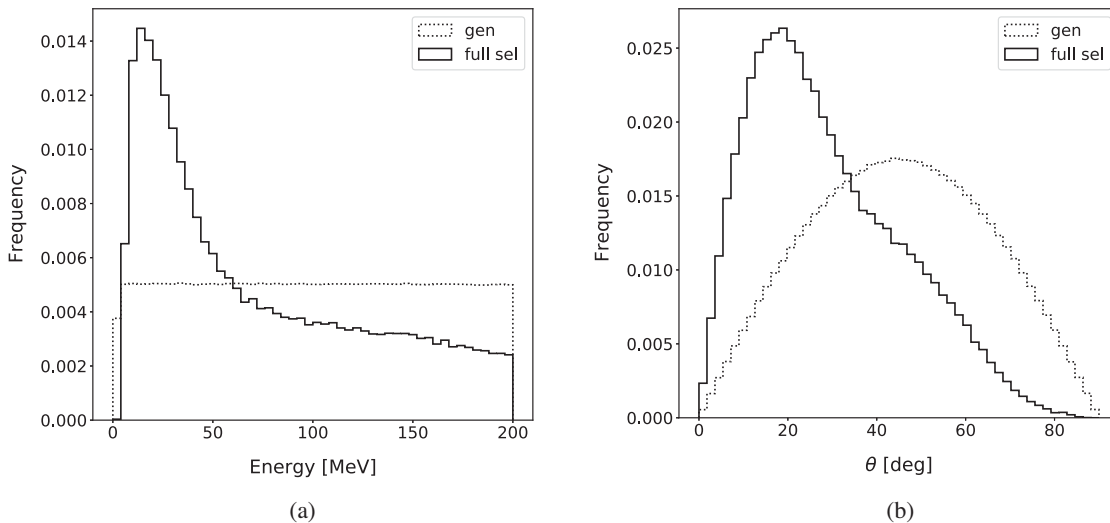Figure 4 shows the longitudinal profile for an electron of 60 MeV and a proton of 150 MeV in ADC. The profile

(a)

(b)

FIG. 3. Normalized distribution of energy (a) and polar angle θ (b) for simulated electrons are shown in dotted black. The solid black distribution represents events traversing the trigger and the first two calorimeter planes without hitting the lateral veto. The distributions are similar for protons.

provides an indication on the type of particle observed by HEPD. In fact, most electrons steadily deposit energy in the planes, whereas contained protons exhibit a Bragg peak [19] or an increasing amount of signal with the number of traversed planes before reaching the LYSO crystals.

The signals collected by the different sub-detectors are used as the input for the FCNNs architecture, since they are
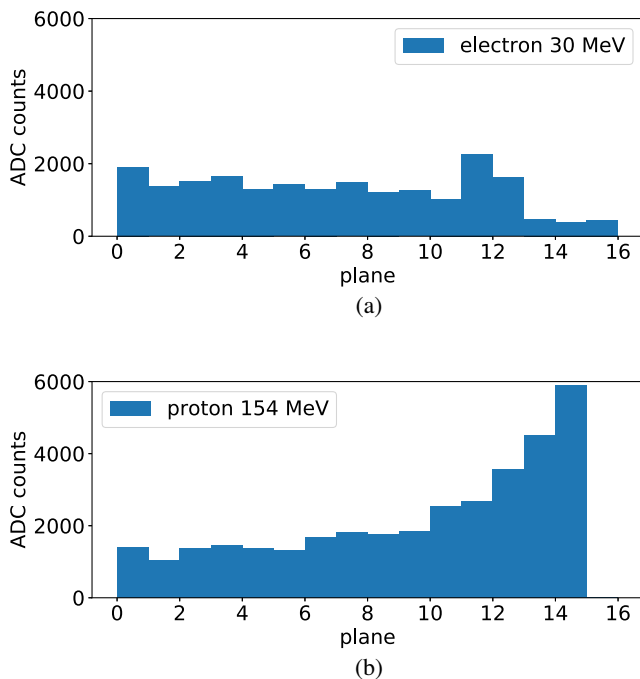

(a)

(b)

FIG. 4. Signal profiles for the simulation of an electron (a) and a proton (b). In the latter case the Bragg peak in the 14th-15th planes is clearly visible. The vertical axis reports the sum of ADC counts from the pair of PMTs coupled to each scintillating tile.

correlated with the type of the incoming particle, its arrival direction and its energy.

These dependencies are well visible in the correlation matrix of Fig. 5, built with events passing the preselection (1). It can be noticed how the angles are correlated with the energy and the signals in the PMTs. The signals on the last PMTs or on the LYSO crystals are anticorrelated with θ, because only vertical tracks are allowed to reach to the deepest objects. For what concerns ϕ or the y coordinate of the impact position on the reference plane, they are strongly correlated with the trigger bars and with the signals from the LYSO crystals, since they carry positional information. It can be seen how the energy is anticorrelated with θ, for the same reason as before, i.e., the more energetic the particle, the deeper the penetration, increasing the chance to fail passing the preselection cut. Moreover, for protons [Fig. 5(b)], the signal from the upstream scintillating tiles anticorrelates with energy: in the kinetic energy range 30–1000 MeV (Lorentz factor $\gamma = 1 \div 0.7$) the energy loss occurs primarily via ionization and higher energy protons release smaller amounts of energy in the tiles.

Types of particles other than protons and electrons have not been included in the training set, because the chain of reconstruction primarily aims at selecting and measuring protons and electrons. The main inefficiency in particle identification concerns m.i.p. and quasi-m.i.p. mimicking electrons. Most of these cases are due to protons of energy $E > 200$ MeV, abundantly included in the training set of protons. After the preselection, the measurement of energy and arrival direction proceeds with different networks for electrons and protons. For electrons, no improvement of performance is expected when positrons and gammas are included for training, the former being practically indistinguishable from electrons and the latter never triggering
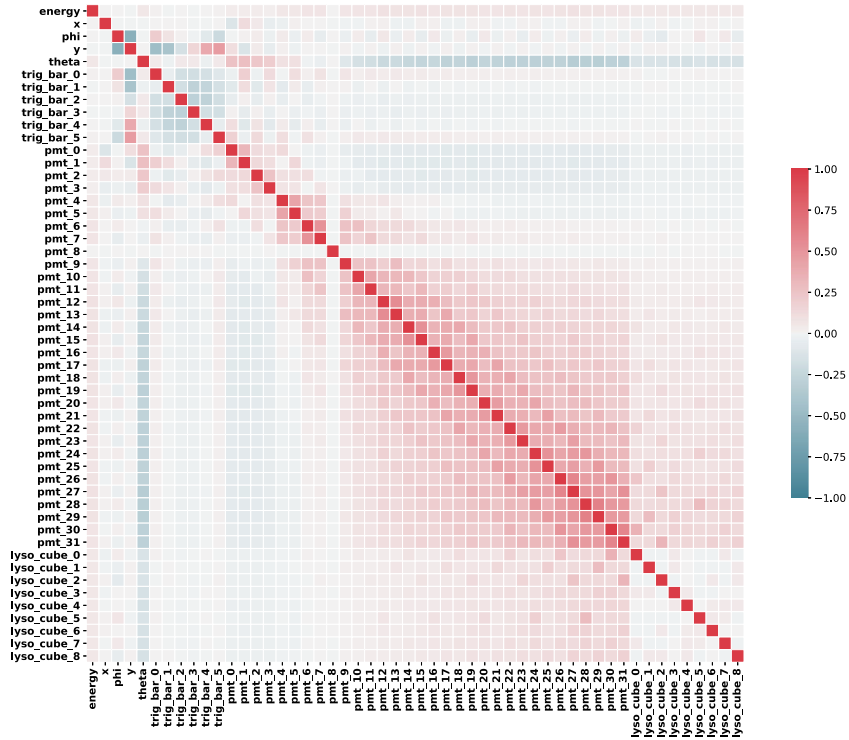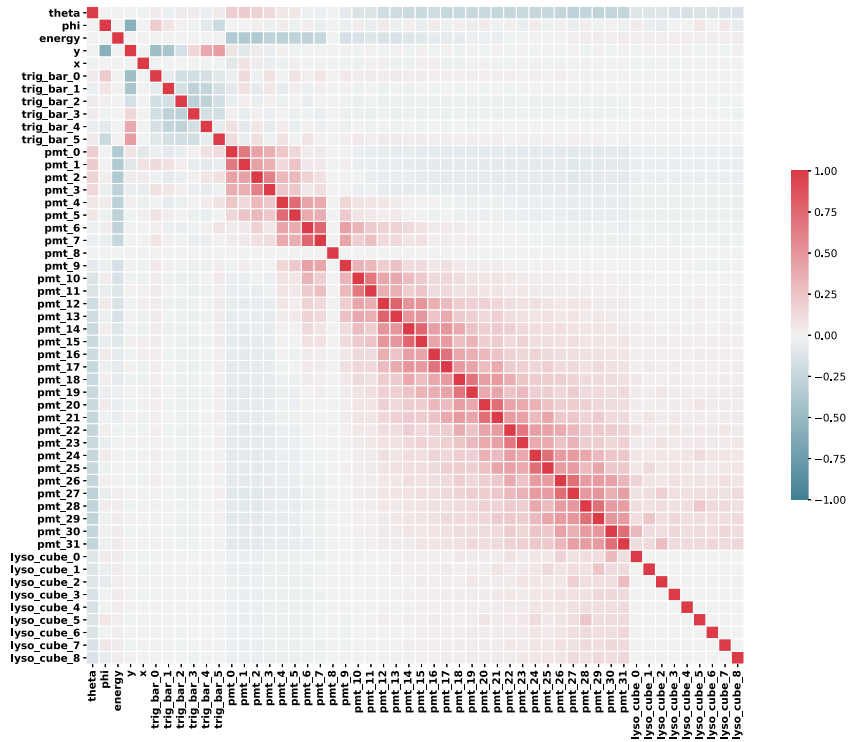
FIG. 5.   Correlation matrices between input and target variables for simulated electrons (a) and protons (b). The color scale represents the linear correlation coefficient between variables.

the acquisition. Concerning the performance on protons, the impact of deuterons, nuclei and isotopes with $Z \geq 2$ have been found statistically negligible, since their estimated integral fluxes are more then two order of magnitudes below the proton one in a usual HEPD orbit. At the time of writing, training data sets are being extended to include protons with kinetic energy up to 10 GeV and helium nuclei up to 1 GeV.

IV. EVENT RECONSTRUCTION BASED ON A DNN

A. Reconstruction chain

The goal of this work is the identification of the particle that traverses the detector, and the measurement of its energy (E_{kin}) and arrival direction (ϕ, θ).

The input available for the algorithms is a set of 52 numbers per event: 32 PMT signals from 16 plastic scintillators and 9 from the LYSO crystals; 6 booleans from the trigger bars and 5 from the veto planes. Signals from plastic scintillators and LYSO crystals are used as they come from data acquisition or Monte Carlo digitization. The digitization of Monte Carlo simulations is consistent with the calibration curves reported in [18] and inherits systematic uncertainties described therein. The calibration at beam tests did not regard trigger bars and veto planes, therefore signals from these subsystems are transformed into booleans before use (1 if the signal is above threshold, 0 otherwise). This device makes networks more robust against statistical and systematic effects of the Monte Carlo digitization.

The input vector is used to train a reconstruction chain based on three fully connected neural networks (FCNNs) [20]: one to identify the type of particle (FCNN$_{pid}$) and two others to predict the energy, the arrival direction and the impact position of the incoming particle. These two networks (FCNN$_{kin}$) are independent of each other and they are tuned for electrons and protons. FCNN$_{pid}$ is trained on a balanced dataset of protons and electrons, while the two FCNN$_{kin}$ are trained only on the targeted particle type. The authors applied FCNN$_{kin}$ on inputs from misidentified particles. Information useful to assess systematic effects on physics analyses was retrieved, but no need for modification of the training dataset emerged.

The reconstruction algorithm starts determining the type of particle using FCNN$_{pid}$. The outcome of the classification is then used to decide which FCNN$_{kin}$ must be used to reconstruct the kinematics of the event, whether that for electrons or for protons.

All FCNNs share the same architecture: they have six layers, i.e., the input layer, four hidden layers and the output layer, with 52, 512, 512, 512, 256 and n nodes respectively. $n = 1$ for FCNN$_{pid}$ and $n = 7$ for FCNN$_{kin}$. All layers have a `Relu` [21] activation function, except the output layer of FCNN$_{pid}$, which uses a `Sigmoid` activation. The `Dropout` [22] layers are placed between the

hidden ones to prevent overfitting. The dropout parameter p, representing the fraction of dropout nodes in each hidden layer, is set to 0.1. A grid along the number of layers and nodes, the hyperparameters and the dropout value has been devised and explored to optimize the architecture of FCNNs.

B. Training procedure

The networks are implemented using a combination of Sklearn [23] and PyTorch [24] libraries that provide useful tools for training and validation in the design and the construction of the models. For the particle type separation, the model is trained minimizing a `BCELoss`[1] [25]. For the prediction of the energy and of the arrival direction (θ, ϕ) the models were trained minimizing the sum of different terms, where all terms are scaled to contribute with the same weight to the total loss.

$$\text{Loss}_{\text{total}} = \text{Loss}_{E_{kin}} + \text{Loss}_{(\theta,\phi)} + \text{Loss}_{\text{control}} \quad (2)$$

The `L1Loss`[2] metric is adopted for the term $\text{Loss}_{E_{kin}}$ [26]. It weights more the core of the distribution than outliers. For the term $\text{Loss}_{(\theta,\phi)}$, a loss representing the angular difference between the true and the predicted direction is used:

$$\Psi = \arccos(\vec{v}_{tr} \cdot \vec{v}_{pr})$$

$$\vec{v}_{tr} = \begin{pmatrix} \sin\theta_{tr}\cos\phi_{tr} \\ \sin\theta_{tr}\sin\phi_{tr} \\ \cos\theta_{tr} \end{pmatrix}, \qquad \vec{v}_{pr} = \begin{pmatrix} \sin\theta_{pr}\cos\phi_{pr} \\ \sin\theta_{pr}\sin\phi_{pr} \\ \cos\theta_{pr} \end{pmatrix} \quad (3)$$

where $\vec{v}_{pr}$ and $\vec{v}_{tr}$ are the unitary vector of the true direction and that of the predicted direction, respectively. In addition to the energy E_{kin} and the arrival direction (θ, ϕ), other control variables are predicted, i.e., the impact position (x, y) on the reference plane, the energy released in the tower of scintillator planes and the energy released in the LYSO crystals, using the metric `L1Loss`. The two auxiliary energies contribute to supervise the model, yet unneeded for the event reconstruction. Conversely, x and y are exploited to enforce a fiducial selection that will be described in the following.

For all FCNNs, 80% of the Monte Carlo (MC) dataset is used for training and validation, while the remaining 20% for test. In order to reduce the number of epochs necessary during the training, a learning rate (LR) scheduler [27] starting from a LR of 10^{-3} and a 2% decrease each epoch is used. Figure 6 reports training and validation losses for FCNN$_{kin}$ and FCNN$_{pid}$. In order to choose the best model during training and to prevent overtraining, a checkpoint

---

[1] $\text{BCELoss} = \sum_{i=0}^{N} y_n \cdot \log x_n + (1 - y_n) \cdot \log(1 - x_n)$
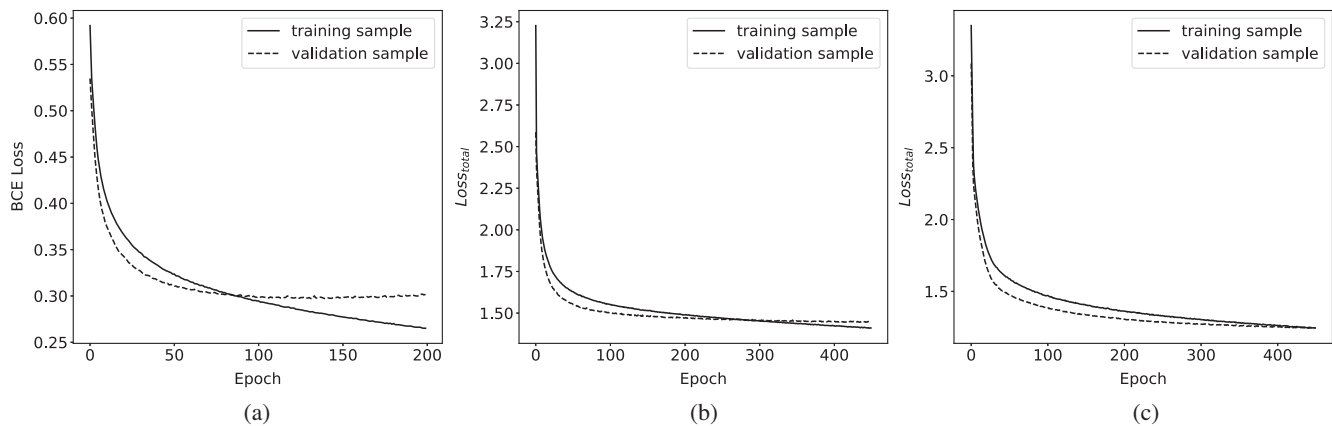[2] $\text{L1Loss} = \sum_{i=0}^{N} |y_{true}^i - y_{pred}^i|$

FIG. 6.   Validation losses for $FCNN_{pid}$ (a), $FCNN_{kin}$ for electrons (b) and $FCNN_{kin}$ for protons (c), represented for training and validation as a function of training epoch.

system, saving the model as long as the validation loss decreases, is employed.

V. RESULTS WITH MONTE CARLO

In this section the performance obtained using the NN based algorithm on MC samples are presented. The performance has been evaluated on a set of simulated events statistically independent of the training dataset, after applying the preselection (1).

A. Particle identification

$FCNN_{pid}$ takes as input the signal of all photomultiplier tubes (PMTs). It uses the information of the PMT signals and the plane-by-plane profile of the signal to distinguish electrons from protons. The output of the network is a particle identification (PID) number, ranging from 0 to 1. We set 0.5 as the threshold above which the particle is identified as an electron, otherwise the particle is tagged as a proton. The efficiency is defined as follows:

$$\text{Eff}_{\text{part}} = \frac{N_{\text{ev}}(\text{TAG} \equiv \text{part})}{N_{\text{ev}}(\text{TRUTH} \equiv \text{part})}$$

where "N_{ev}" stands for number of events, "part" indicates either electron or proton, "TAG" is the reconstructed particle type and "TRUTH" is the MC true particle type. The efficiency is shown as a function of the primary kinetic energy in Fig. 7.

Tagging efficiency for electrons exceeds 99% from the threshold up to kinetic energies as large as $E_{\text{u.c.}}^{\text{elec}} = 30$ MeV. For protons, it never drops below 95% up to $E_{\text{u.c.}}^{\text{prot}} = 150$ MeV, being stable at 98% above threshold. $E_{\text{u.c.}}^{\text{elec}}$ and $E_{\text{u.c.}}^{\text{prot}}$ are the energy limits within which the energy release mostly takes place in the upper calorimeter. This condition favors particle identification because the event topology of a proton exhibiting a Bragg peak is easily distinguished from the smooth profile of an electron (see

Fig. 4). For energies larger than $E_{\text{u.c.}}^{\text{prot}}$, the tagging efficiency for protons drops because the Bragg peak is not so well localized in LYSO crystals as in plastic scintillator tiles. For kinetic energies larger than 500 MeV (Lorentz boost $\gamma \gtrsim 1.6$), protons behave like quasi-m.i.p., making it more likely to tag them as electrons. The same effect is visible in the smooth decrease of the efficiency of electrons after $E_{\text{u.c.}}^{\text{elec}}$: the prediction of $FCNN_{pid}$ worsens due to $E \gg E_{\text{u.c.}}^{\text{elec}}$ electrons being possibly confused with $E \gg E_{\text{u.c.}}^{\text{prot}}$ protons.

B. Fiducial selection for the reconstruction of kinetic quantities

As in other approaches to the event reconstruction, performance depends on the selection applied to data. Given the size of the HEPD effective area, almost two orders of magnitude larger than competitors and predecessors [28,29], a tight selection on the predicted impact
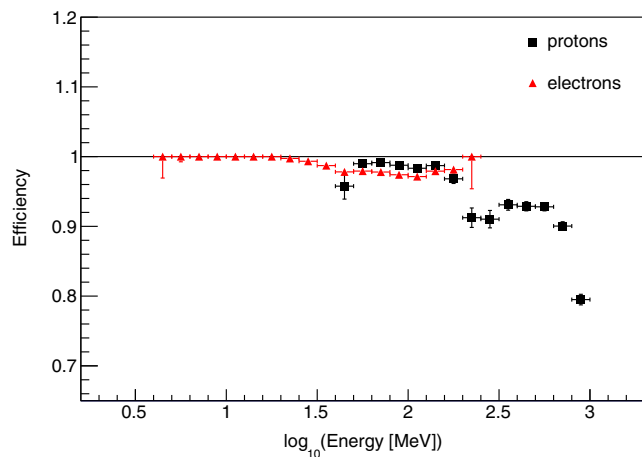

FIG. 7.   DL-based particle identification efficiency as a function of the kinetic energy of primary electrons (red triangles) and protons (black squares). Error bars represent statistical uncertainty.
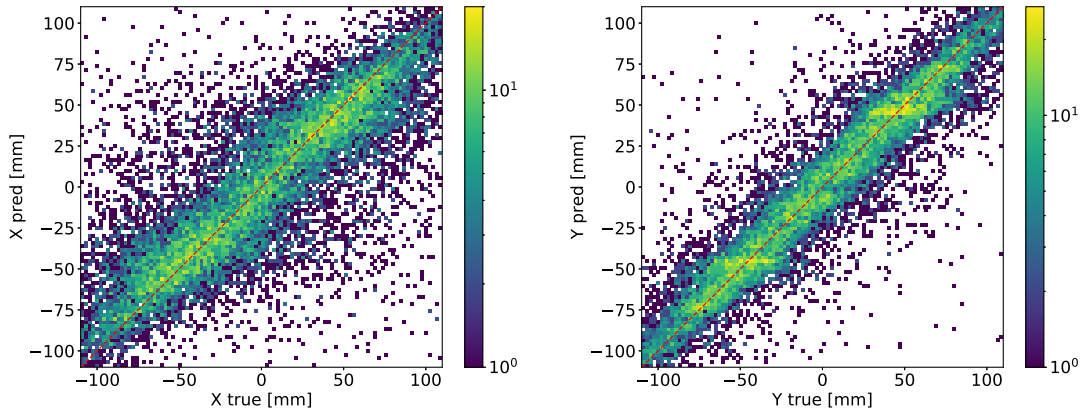
FIG. 8.   Predicted versus true position for protons after preselection (1). The color scale specifies the number of events per bin. The red line is the bisector.

position of the particle was envisaged, i.e., the control variables x and y. The performance of the prediction on these variables is shown for protons in Fig. 8. The y coordinate is better determined than x because trigger bars are deployed along the y direction.

Before the launch, beam tests with protons and electrons were carried out with particles entering HEPD at different points. The positions with the largest fraction of statistics are $P_1(x, y) = (-17.5, -18)$ mm and $P_2(x, y) = (17.5, 18)$ mm on the reference plane, where the axes origin lays on the axis of the apparatus. P_1 and P_2 are symmetric with respect to the instrument axis. The MC digitization and the calibration of the HEPD response have been carried out using data collected during beam tests at points P_1 and P_2. To achieve full control of the estimate of the performance and to suppress effects due to data/MC discrepancies, the following fiducial selection was applied, consisting of the union of two squares of 15 mm side, centered at points P_1 and P_2:

$$(x, y) \in \text{square}(P_1, 15 \text{ mm}) \cup \text{square}(P_2, 15 \text{ mm}) \quad (4)$$

C. Kinetic energy reconstruction

The performance of FCNN_{kin} is different for protons and electrons, according to different mechanisms of the energy release in scintillators.

Figure 9 reports the performance of the energy estimation by FCNN_{kin} for protons and electrons after the fiducial selection. Good performance is obtained for protons even above the energy limit of containment, corresponding to the design value of 200 MeV. This result is due to FCNN_{kin} being sensitive to the increasing energy loss per unit length that protons suffer while reaching the Bragg peak, even if the peak itself is not observed. Saturation effects, related to the energy range chosen for training, are barely visible. Electrons energy is reconstructed with poorer resolution. As for protons, the lowest energy interval mostly suffer from unobserved energy deposited in passive materials. But
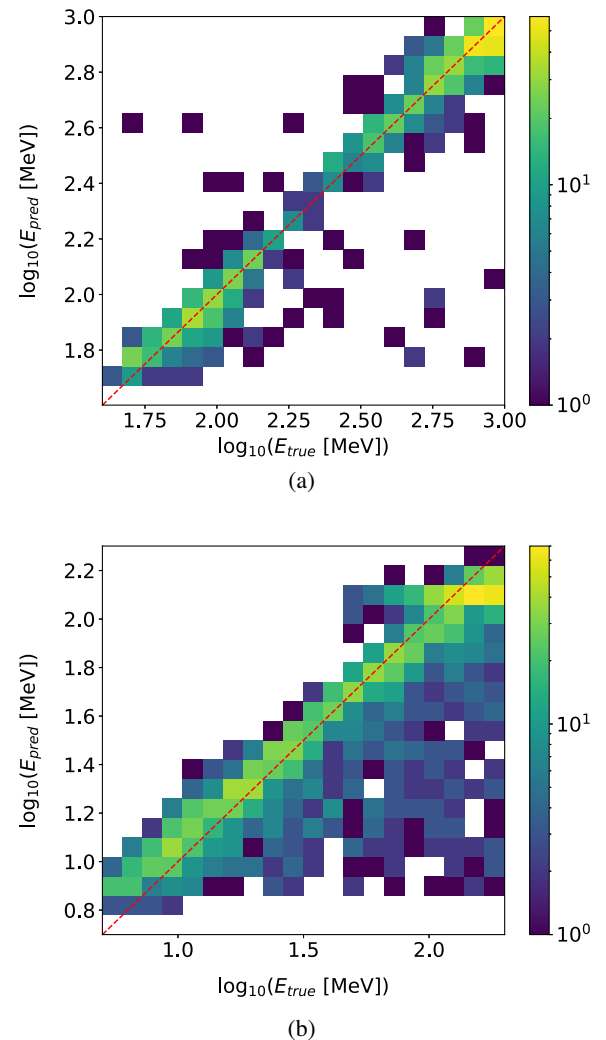

(a)

(b)

FIG. 9.   Energy predicted by the DL-based algorithms as a function of true energy for protons (a) and electrons (b). The color scale represents the number of events in each bin. The red line is the bisector.

relevant amounts of energy can leave HEPD undetected also for electrons of higher energy, due to Bremsstrahlung. The predicted energy for electrons shows clustering at about 150 MeV. This effect is due to events with energy limited by containment (100 MeV) and maximum MC energy (200 MeV): their energy releases do not sufficiently differ from each other to allow FCNN$_{\text{kin}}$ to make accurate predictions. The mid-value of 150 MeV is often provided as output, as it minimizes the loss function. Such an effect has little or no impact on data analysis, because no electrons with energy larger than 100 MeV are usually considered. Bremsstrahlung photons are copiously emitted by electrons from few tens of MeV on, often leaving the detector unobserved. This radiation is the main contribution to the spread of the energy measurement released by the primary particles. When the FCNN has to predict the kinetic energy, the energy interval over which the prediction is allowed, plays a role, biasing the result by excess.

To account for this effect, the following correction has been derived with a linear fit in the containment energy range 0–100 MeV (0–200 MeV) for electrons (protons):

$$E_{\text{true}}^{\text{ele}} = 1.01(\pm 0.04)_{\text{stat}} \cdot E_{\text{pred}}^{\text{ele}} - 2.03(\pm 0.93)_{\text{stat}} \quad (5)$$

$$E_{\text{true}}^{\text{prot}} = 0.99(\pm 0.03)_{\text{stat}} \cdot E_{\text{pred}}^{\text{prot}} - 1.53(\pm 3.23)_{\text{stat}} \quad (6)$$

where all variables and numbers are in MeV.

The energy resolution after corrections (5) and (6) is reported in Fig. 10, in the energy ranges targeted by HEPD for electrons and protons. It is obtained by fitting the distribution of the difference $E_{\text{pred}} - E_{\text{true}}$ to a standard gaussian and normalizing σ to the predicted energy. Averaged over of the full solid angle accessible to HEPD, the resolution improves with the energy, reaching 10% for 80 MeV electrons and 6% for 150 MeV protons.
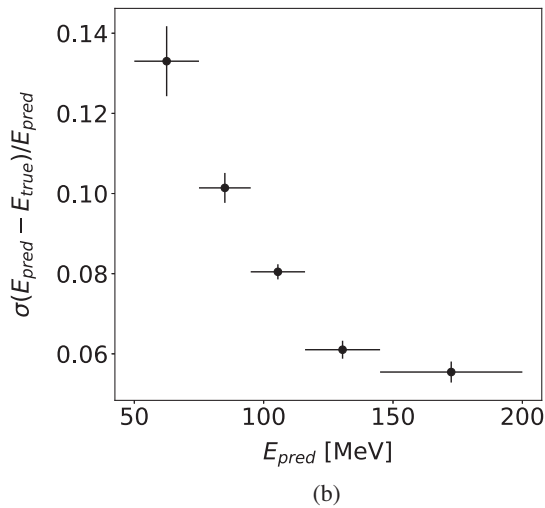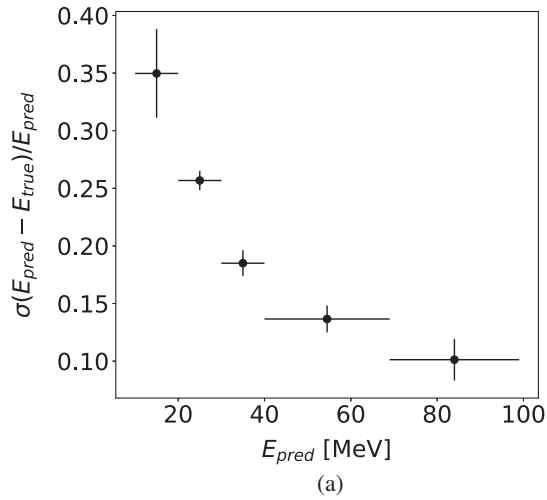

FIG. 10. Energy resolution of FCNN$_{\text{kin}}$ as a function of the predicted energy for electrons (a) and protons (b). Error bars represents the statistical uncertainty. See the text for details.
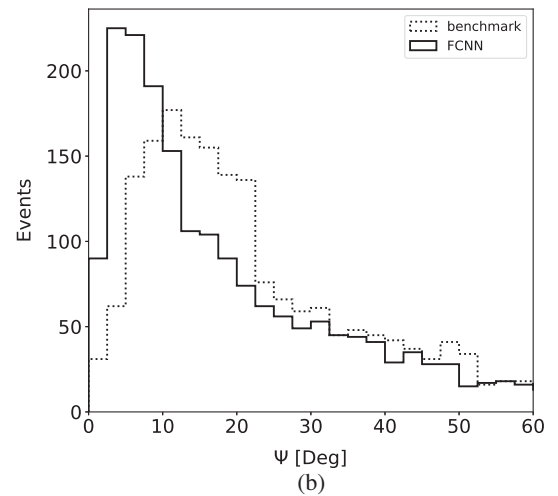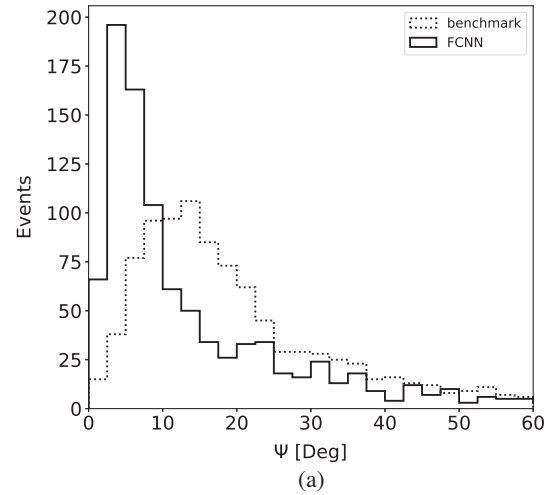

FIG. 11. Angular difference Ψ between the predicted arrival direction and the true one is shown for protons (a) and electrons (b).

TABLE I. Angular reconstruction performances are shown for electrons and protons. The "vertical" case represent the case in which the vertical pointing is used as the predicted direction.

Particle type	Method	Mode	FWHM	90%
Protons	FCNN	3.7	7.5	32
Protons	Benchmark	13.7	17.5	38
Electrons	FCNN	3.7	10.0	40
Electrons	Benchmark	11.2	17.5	43

D. Arrival direction reconstruction

The set of detectors considered in this work is not designed to be sensitive to the arrival direction of the impinging particle. It is quite challenging to estimate θ and ϕ on the basis of the light produced in scintillators, scattered all around the sensitive volumes (the attenuation length for HEPD plastic scintillators is 380 cm [30]) and then collected by a sparsified readout of PMTs. The DL-based reconstruction strategy exploits the following information: (i) correlations between PMTs responses coupled with the same scintillator; (ii) coarse positional information from trigger bars and LYSO crystals; (iii) anticorrelation between polar angle and kinetic energy, as shown by the correlation matrices in Fig. 5. Without $FCNN_{kin}$, the best estimation of (θ, ϕ) for each event would be (0,0) (the axis of the apparatus), because no information on the arrival direction is directly attainable from plastic and LYSO crystal scintillators. $(\theta, \phi) = (0, 0)$ is the average expected arrival direction of particles and it constitutes the natural benchmark to evaluate the performance of $FCNN_{kin}$. The distribution of Ψ (see definition (3) in Sec. IV B) for electrons and protons passing selection (4) is shown in Fig. 11, together with the benchmark. After averaging out over the whole energy range, the distributions show that $FCNN_{kin}$ provides remarkable gains over the benchmark. Three figures of merit—mode, full width at half maximum (FWHM) and 90% quantile of Ψ distributions—are

reported in Table I, quantifying the average performance of $FCNN_{kin}$ in reconstructing the arrival direction. As expected, the benchmark gives the same result for electrons and protons, whereas $FCNN_{kin}$ performs better for protons than for electrons, as already seen and motivated for the prediction of kinetic energy.

VI. RESULTS WITH DATA

This section illustrates the performance of $FCNN_{kin}$ applied to real data, collected during beam tests and in flight.

A. Beam-test data

Before launch, a campaign of tests was carried out to calibrate the detector response to charged particles [18]. In October 2016 the detector was irradiated with electrons at the INFN Beam Test Facility (BTF) in Frascati. In November 2016 HEPD was tested with a proton beam delivered at the Trento APSS Proton Therapy Center. In the context of this work, these tests are used to validate the FCNNs on data for which energy and arrival direction are known. Whenever possible, the $FCNN_{pid}$ network was applied to beam test events. The tagging efficiency found with Monte Carlo simulation and described in Sec. V has been confirmed for protons, 98% of which have been correctly tagged in the energy range 37–228 MeV. For electrons in the energy range 30–90 MeV, the tagging efficiency has been found to be 95%. Figure 12 illustrates the performance of $FCNN_{kin}$ for electrons of energy $E_{beam} = 60$ MeV and arrival direction $(\theta, \phi) = (0, 0)$, impacting on positions P_1 and P_2 (19530 events in total, 4728 after selections). To be included in the performance estimation, the events must pass the selection (4) and be tagged as electrons. The predicted energy is compatible with the beam energy declared by the facility and the measured resolution $(7.9 \pm 0.3\%)$ is compatible with the value expected for an isotropic distribution of arrival directions (12%, see Fig. 10). This similarity is confirmed
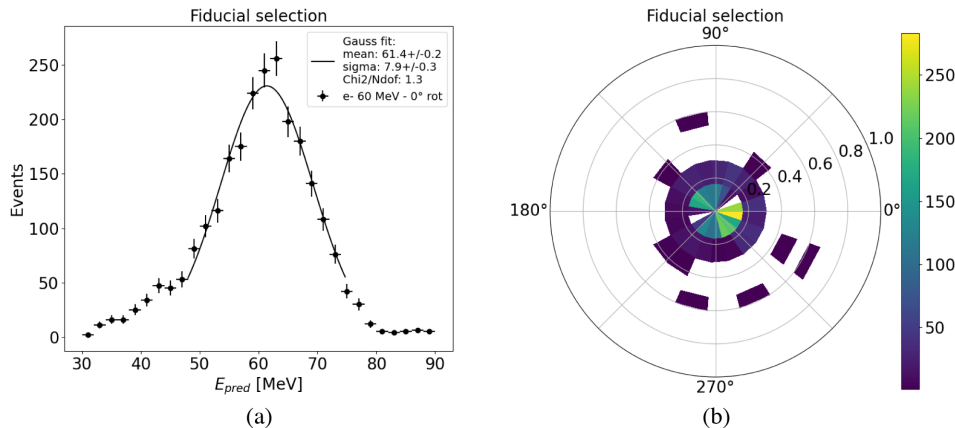

FIG. 12. Kinetic energy (a) and arrival direction (b) predictions from $FCNN_{kin}$ for beamed electrons ($E_{beam} = 60$ MeV and $(\theta, \phi) = (0°, 0°)$). The color scale specifies the number of events.

by MC simulations of isotropic and beamed fluxes and is due to multiple Coulomb scattering [31] that electrons undergo at such low energy. The arrival direction is also efficiently predicted: most events are reconstructed within the circle $\sin\theta = 0.2$, i.e., with polar angle less than 12°.

Figure 13 illustrates the performance of $FCNN_{kin}$ for protons. Also in this case, only events passing the selection (4) and tagged as protons are included. Two cases are reported: (i) energy $E_{beam} = 154$ MeV and arrival direction $(\theta, \phi) = (0, 0)$ (30001 events in total, 10003 after selections); (ii) energy $E_{beam} = 70$ MeV and arrival direction $(\theta, \phi) = (15°, 270°)$ (21590 events in total, 13323 after selections). Protons impacted on HEPD on positions P_1 and P_2. In the second case the predicted energy is 3 MeV lower than E_{beam}. This prediction is well compatible with MC expectations, showing that 70 MeV protons lost 3 MeV in air before reaching HEPD in the APSS experimental hall. In both cases, the energy resolution turns out to be better than what has been found for the average result shown in Fig. 10. This effect is correctly reproduced by MC simulations of proton beam tests, where multiple Coulomb scattering does not dilute the dependence of the energy resolution on the polar angle, as it does for

electrons. Indeed, the performance of $FCNN_{kin}$ 70 MeV protons beamed at $\theta = 0°$ is better than the average one obtained for the same energy but θ randomly sampled in 0–90°. Concerning the arrival direction, it can be seen how $FCNN_{kin}$ correctly predicts $(\sin\theta, \phi) = (0.3, 270°)$, corresponding to $(\theta, \phi) = (15°, 270°)$.

B. Flight data

On the 2nd of February 2018 the CSES-01 mission started with the launch of the satellite from China. HEPD and all other payloads have been smoothly operated since then, and scientific data are regularly transmitted to the ground segment [32,33]. The reconstruction chain described in this work is included in data processing and executed on all events collected by HEPD. In this section, the performance of $FCNN_{kin}$ is described for data collected in flight on August 1°, 2018 (first day of post-commissioning operations). Out of 5903533 events, 35834 pass the selections (1) and (4).

Since MC events have been generated with flat energy spectrum, they have to be reweighted before comparing them with flight data. Nonetheless, HEPD is triggered by particles of different types and different populations along
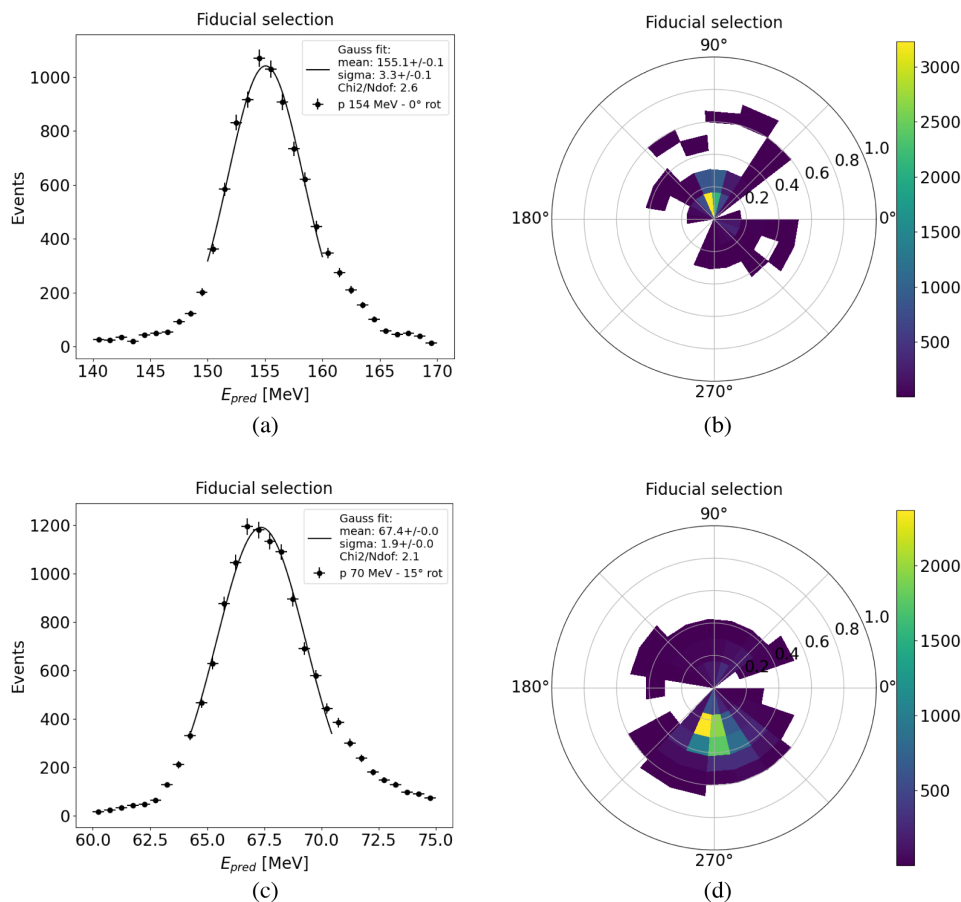

FIG. 13. Kinetic energy and arrival direction predictions from $FCNN_{kin}$ for beamed protons. In plots (a) and (b) protons have $E_{beam} = 154$ MeV and $(\theta, \phi) = (0°, 0°)$. In plots (c) and (d) protons have $E_{beam} = 70$ MeV and $(\theta, \phi) = (15°, 270°)$. The color scale specifies the number of events.
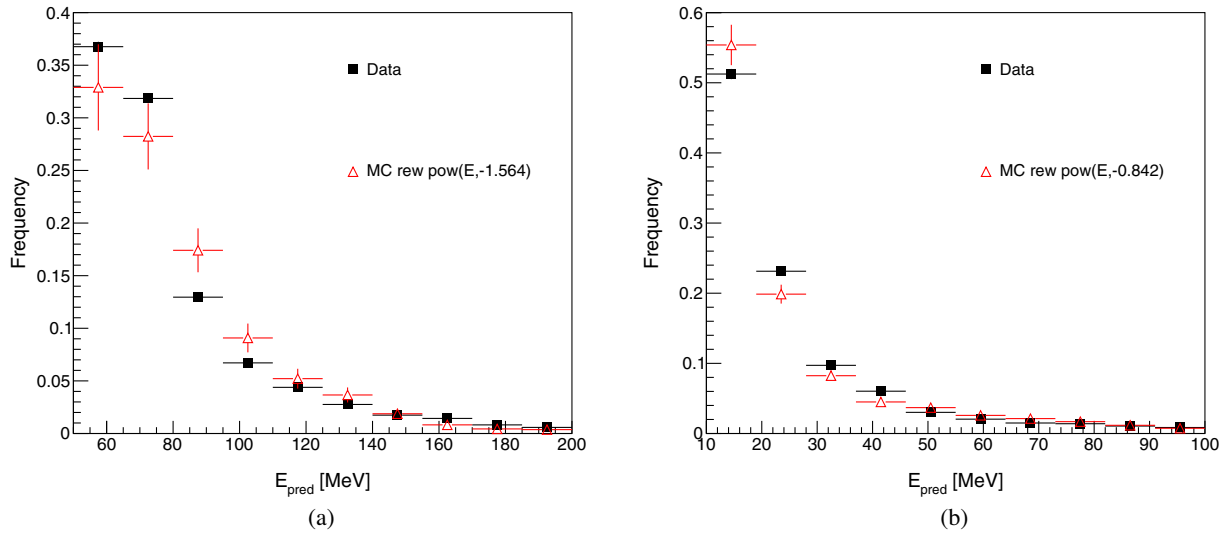
FIG. 14.   Distribution of reconstructed energy for protons (a) and electrons (b). Monte Carlo events (red triangles) are reweighted as described in the text. Error bars represent statistical uncertainty.

its orbit: trapped electrons and positrons, trapped protons, galactic electrons, galactic protons, electrons and protons from the South Atlantic Anomaly [34]. All of these populations have different energy spectra and to separate each contribution goes beyond the scope of this work. The authors opted for reweighting MC events on the basis of their true energy, using power laws:

$$w_{\mathrm{ele}} = E_{\mathrm{true}}^{\alpha_{\mathrm{ele}}}, \qquad w_{\mathrm{prot}} = E_{\mathrm{true}}^{\alpha_{\mathrm{prot}}}$$

The spectral indices α_{ele} and α_{prot} are obtained by a χ^2 minimization procedure, fitting the MC energy distribution to flight data tagged as electrons and protons respectively. The best indices were found to be $\alpha_{\mathrm{ele}} = -0.842$ and $\alpha_{\mathrm{prot}} = -1.564$. Figure 14 shows that the power law

hypothesis guarantees fair agreement between flight and MC data on the reconstructed energy distribution.

Since all particle populations enter the data energy distribution, the best-fit values of α_{ele} and α_{prot} have to be considered as effective indices, useful to compare the performance of $FCNN_{\mathrm{kin}}$ on flight and MC data for what concerns the prediction of the arrival distribution.

In fact, good agreement can be observed in Fig. 15, where the distribution of the polar angle θ is represented for flight and MC data, for both electrons and protons. Within uncertainties, the prediction of $FCNN_{\mathrm{kin}}$ on MC data matches that on flight data, up to fairly large values (40°). The authors observe that no pointlike sources are available to test the performance on the estimation of the arrival direction, neither patterns of (θ, ϕ) due to magnetic
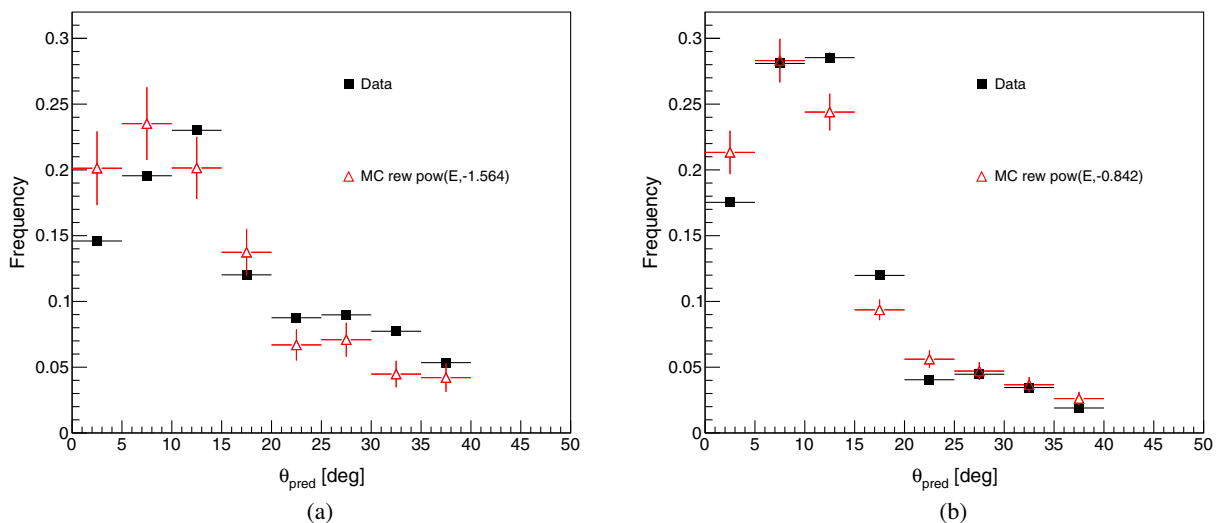

FIG. 15.   Distribution of reconstructed polar angle θ for protons (a) and electrons (b). Monte Carlo events (red triangles) are reweighted as described in the text. Error bars represent statistical uncertainty.

field gradients or time variations of particle fluxes can be easily exploited. Indeed, averaging out these dependencies on time and geomagnetic latitude is the main source of systematic uncertainty, that anyway is less than the angular resolution mentioned in Table I. The result shown in Fig. 15 is the simplest and most powerful check that the reconstruction of the arrival direction of electrons and protons by FCNN$_{kin}$ is reliable.

VII. CONCLUSIONS

The reconstruction of events collected by the High Energy Particle Detector on board the CSES satellite has been based on a deep learning approach. Fully connected neural networks allow to efficiently tag protons and electrons, and to reconstruct with good accuracy the energy of incoming particles. Remarkably, it has been demonstrated that:

(i) the arrival direction can be reasonably inferred, just using information from detectors insensitive to the impact position;

(ii) the networks efficiently reconstruct the energy also for noncontained particles, a result difficult to achieve with methods based on analytical extrapolation of the energy deposit versus the traversed material thickness.

Results from Monte Carlo simulations have been validated with data from beam tests and flight operations.

[1] ATLAS Collaboration, J. Instrum. **11,** P04008 (2016).
[2] CMS Collaboration, Comput. Software Big Sci. **4,** 10 (2020).
[3] A. Khan, A. Chefranov, and H. Demirel, Neurocomputing; Variable Star Bulletin **440,** 111 (2021).
[4] G. Ambrosi, S. Bartocci, L. Basara *et al.*, Sci. China Technol. Sci. **61,** 643 (2018).
[5] X. Shen, X. Zhang, S. Yuan *et al.*, Sci. China Technol. Sci. **61,** 634 (2018).
[6] P. Picozza, R. Battiston, G. Ambrosi *et al.*, Astrophys. J. Suppl. Ser. **243,** 16 (2019).
[7] M. Martucci, R. Sparvoli *et al.*, Appl. Sci. **11,** 3465 (2021).
[8] F. Palma, A. Sotgiu, A. Parmentier, M. Martucci, M. Piersanti, S. Bartocci, R. Battiston, W. J. Burger, D. Campana, L. Carfora *et al.*, Appl. Sci. **11,** 5680 (2021).
[9] S. Bartocci, R. Battiston *et al.*, Astrophys. J. **901,** 8 (2020).
[10] R. Battiston and V. Vitale, Nucl. Phys. B, Proc. Suppl. **243–244,** 249 (2013), Proceedings of the IV International Conference on Particle and Fundamental Physics in Space.
[11] V. Sgrigna, A. Buzzi, L. Conti, P. Picozza, C. Stagni, and D. Zilpimiani, Tectonophysics **431,** 153 (2007), Mechanical and electromagnetic phenomena accompanying preseismic deformation: From laboratory to geophysical scale.
[12] P. Picozza, L. Conti, and A. Sotgiu, Front. Earth Sci. **9,** 578 (2021).
[13] C. Fidani, R. Battiston, W. Burger, and L. Conti, Int. J. Rem. Sens. **33,** 4796 (2012).
[14] Y. Kopytenko, T. Matiashvili, P. Voronov, E. Kopytenko, and O. Molchanov, Phys. Earth Planet. Inter. **77,** 85 (1993).
[15] A. C. Fraser-Smith, P. R. McGill, R. A. Helliwell, and O. G. Villard Jr., Geophys. Res. Lett. **21,** 2195 (1994).
[16] S. Y. Aleksandrin, A. M. Galper, L. A. Grishantzeva, S. V. Koldashov, L. V. Maslennikov, A. M. Murashov, P. Picozza, V. Sgrigna, and S. A. Voronov, Ann. Geophys. **21,** 597 (2003).
[17] S. Agostinelli *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **506,** 250 (2003).
[18] G. Ambrosi *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **974,** 164170 (2020).
[19] P. W. Bragg M. A., London, Edinburgh, Dublin Philos. Mag. J. Sci. **8,** 719 (1904).
[20] A. Zhang, Z. C. Lipton, M. Li, and A. J. Smola, arXiv:2106.11342.
[21] V. Nair and G. E. Hinton, in *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10 (OmniPress, Madison, WI, USA, 2010), p. 807814.
[22] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, J. Mach. Learn. Res. **15,** 1929 (2014), http://jmlr.org/papers/v15/srivastava14a.html.
[23] L. Buitinck *et al.*, arXiv:1309.0238.
[24] A. Paszke *et al.*, in *Advances in Neural Information Processing Systems 32*, edited by H. Wallach, H. Larochelle, A. Beygelzimer, F. dAlché-Buc, E. Fox, and R. Garnett (Curran Associates, Inc., New York, 2019), pp. 8024–8035.
[25] PyTorch implementation of BCE, https://pytorch.org/docs/stable/generated/torch.nn.BCELoss.html (2021) (Accessed: 2021-08-04).

[26] PyTorch implementation of L1Loss, https://pytorch.org/docs/stable/generated/torch.nn.L1Loss.html (2021) (Accessed: 2021-08-04).

[27] StepLR scheduler documentation page, https://pytorch.org/docs/stable/generated/torch.optim.lr_scheduler.StepLR.html (2021) (Accessed: 2021-08-04).

[28] J. Sauvaud, T. Moreau, R. Maggiolo, J.-P. Treilhou, C. Jacquey, A. Cros, J. Coutelier, J. Rouzaud, E. Penou, and M. Gangloff, Planet. Space Sci. **54**, 502 (2006), First Results of the DEMETER Micro-Satellite.

[29] W. Cook, A. Cummings, J. Cummings, T. Garrard, B. Kecman, R. Mewaldt, R. Selesnick, E. Stone, D. Baker, T. von Rosenvinge, J. Blake, and L. Callis, IEEE Trans. Geosci. Remote Sens. **31**, 565 (1993), cited By 102.

[30] General purpose EJ-200, EJ-204, EJ-208, EJ-212, https://eljentechnology.com/products/plastic-scintillators/ej-200-ej-204-ej-208-ej-212 (2021) (Accessed: 2021-08-04).

[31] W. T. Scott, Rev. Mod. Phys. **35**, 231 (1963).

[32] A. Sotgiu, C. De Donato, C. Fornaro, S. Tassa, M. Scannavini, D. Iannaccio, G. Ambrosi, S. Bartocci, L. Basara, R. Battiston *et al.*, Software Pract. Ex **51**, 1459 (2021).

[33] G. Ambrosi, S. Bartocci, L. Basara, R. Battiston, W. Burger, D. Campana, M. Caprai, L. Carfora, G. Castellini, P. Cipollone *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **1013**, 165639 (2021).

[34] J. Heirtzler, J. Atmos. Sol. Terr. Phys. **64**, 1701 (2002), Space Weather Effects on Technological Systems.

[35] www.leos.ac.cn/ (Accessed on 1 May 2021).